Journal of
Marine Science and Engineering

MDPI

Article

Deep Learning Convolutional Neural Network Applying for the Arctic Acoustic Tomography Current Inversion Accuracy Improvement

Kangkang Jin [1], Jian Xu [1,*], Zichen Wang [1], Can Lu [1], Long Fan [2], Zhongzheng Li [2] and Jiaxin Zhou [2]

[1] School of Marine Science and Technology, Tianjin University, Tianjin 300072, China; jin20212021@163.com (K.J.); wangzichen@tju.edu.cn (Z.W.); lucan@tju.edu.cn (C.L.)
[2] Navy Research Academy, Tianjin 300451, China; fanlong_tianjin@163.com (L.F.); hailangxiou@163.com (Z.L.); JiaXinZhou22021@163.com (J.Z.)
* Correspondence: jian.xu@tju.edu.cn

Abstract: Warm current has a strong impact on the melting of sea ice, so clarifying the current features plays a very important role in the Arctic sea ice coverage forecasting study field. Currently, Arctic acoustic tomography is the only feasible method for the large-range current measurement under the Arctic sea ice. Furthermore, affected by the high latitudes Coriolis force, small-scale variability greatly affects the accuracy of Arctic acoustic tomography. However, small-scale variability could not be measured by empirical parameters and resolved by Regularized Least Squares (RLS) in the inverse problem of Arctic acoustic tomography. In this paper, the convolutional neural network (CNN) is proposed to enhance the prediction accuracy in the Arctic, and especially, Gaussian noise is added to reflect the disturbance of the Arctic environment. First, we use the finite element method to build the background ocean model. Then, the deep learning CNN method constructs the non-linear mapping relationship between the acoustic data and the corresponding flow velocity. Finally, the simulation result shows that the deep learning convolutional neural network method being applied to Arctic acoustic tomography could achieve 45.87% accurate improvement than the common RLS method in the current inversion.

Keywords: ocean acoustic tomography; velocity field; RLS; convolutional neural network; acoustic inverse problem

check for updates

Citation: Jin, K.; Xu, J.; Wang, Z.; Lu, C.; Fan, L.; Li, Z.; Zhou, J. Deep Learning Convolutional Neural Network Applying for the Arctic Acoustic Tomography Current Inversion Accuracy Improvement. *J. Mar. Sci. Eng.* **2021**, *9*, 755. https://doi.org/10.3390/jmse9070755

Academic Editor: Mike Meylan

Received: 12 June 2021
Accepted: 4 July 2021
Published: 8 July 2021

1. Introduction

Computer models suggest that Arctic sea ice is a core climate indicator [1]. According to statistics, in September 2020, the Arctic sea ice extent fell below 4 million square kilometers for the second time, reaching the 2020 minimum. The rapid reduction in Arctic sea ice has accelerated climate warming, affecting the Arctic region and even the global ecosystem and human activities. The subglacial warm current has a strong influence [2] on the distribution and melting of sea ice. Therefore, understanding the flow velocity under the Arctic sea ice is of great significance for understanding the ice–sea–air interaction process, especially for forecasting the Arctic climate and environment.

Currently, limited by the large area of Arctic sea ice, sub-ice acoustic tomography [3] is the only method that can measure the flow field characteristics of the Arctic sea in a large range. In 2000, Skarsoulis [4] introduced the method of matched peak (MPT) inversion to realize the automatic analysis of propagation time values by utilizing the linearization relationship between sound velocity and arrival time. Based on the ray group method, Taniguchi et al. used display solution and regularization inversion [5] to estimate the velocity and verified that a smoother solution could be obtained through regularization. Through least square fitting, harmonic analysis (HA) can be used to determine the amplitude and phase delay of each tidal component, but the least square method will overfit the

non-tidal component and is not suitable for the analysis of non-stationary data [6], so it is difficult to predict turbulent flows. Traditional acoustic tomography algorithm models are often applied in open seas. To improve the accuracy of the current measurement, the model needs to be uniformly linearized during calculation. However, in high-latitude Arctic regions, small-scale ocean phenomena such as internal waves and turbulence affected by Coriolis force [7,8] cause the uneven ocean environment. The flow measurement interference caused by this inhomogeneity cannot be characterized by fixed parameters in sub-ice acoustic tomography calculations. The use of conventional acoustic tomography least squares and other methods tend to produce over-fitting or noise sensitivity problems [9], and the robustness is very poor, which affects the accuracy of acoustic tomography.

Convolutional neural network (abbreviated as CNN) [10] is mainly realized by linear convolution and nonlinear activation function, especially in the inversion of [11–13] nonlinear problems. In this paper, we use the finite element method to construct the sound field and marine environmental parameters. We aim to use CNN to directly analyze the original acoustic data, establish the nonlinear mapping relationship between the acoustic data and the corresponding flow velocity to realize the accurate inversion of the flow field under the ice and solve the problem of the calculation accuracy of the acoustic tomography caused by the nonlinear small-scale interference under the Arctic ice. The simulation results confirm the feasibility and accuracy of this approach.

The main structure of this paper is as follows. First, in Section 2, basic methods and formulas of arctic acoustic tomography are described. Secondly, in Section 3, the limitation of regularized least squares to solve the problem of accuracy improvement is analyzed, and the CNN method is proposed. Then, in Section 4, a simulation experiment based on the finite element is performed, and the performance of the above two algorithms is compared. Finally, conclusions are given in Section 5.

2. Acoustic Tomography Reciprocal Transmission and Flow Field Calculation

2.1. Acoustic Tomography Reciprocal Transmission

The reciprocal transmission process of acoustic tomography is to obtain marine environmental parameters such as temperature currents in the observed sea area by measuring the propagation time of acoustic signals. The acquisition of acoustic signal propagation time depends on ocean acoustic tomography, and the main methods include ray travel time tomography, normal wave phase tomography, peak matching tomography, and matching field tomography. In this study, the most classic ray travel time [14] tomography is used to construct a cost function by accurately measuring the propagation time of acoustic signals placed between the observation stations around the observation sea area to realize the reconstruction of the flow field under the ice. The schematic diagram of acoustic tomography reciprocal transmission flow measurement is shown in Figure 1.

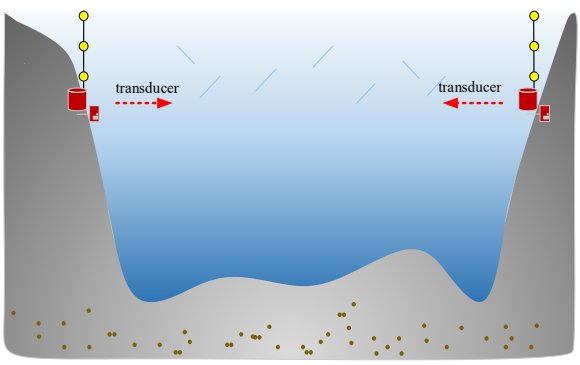

Figure 1. Diagram of reciprocity transmission of acoustic tomography.

The basic theory of ray travel time tomography is ray acoustics, and the ray theory is derived from the wave equation. Considering the propagation law of sound waves in water, the wave equation is described as:

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0 \tag{1}$$

The function $\varphi(x,y,z)$ is introduced, and the formal solution $\mathrm{p}(x,y,z,t) = A(x,y,z)e^{j[\omega t - k_0\varphi(x,y,z)]}$ of the equation is substituted into the wave equation. When $\frac{\nabla^2 A}{A} \ll k^2$ (that is, under the condition of high frequency), the following can be obtained:

$$(\nabla\varphi)^2 = \left(\frac{c_0}{c}\right)^2 = n^2(x,y,z) \tag{2}$$

$$\nabla\cdot\left(A^2\nabla\varphi\right) = 0 \tag{3}$$

Equation (2) is called a function equation, and Equation (3) is an intensity equation, which are two basic equations in ray acoustics. The function equation not only gives the direction of the acoustic ray, but also provides the track and propagation time of the sound line. According to the function equation:

$$\mathrm{n} = \sqrt{(\nabla\varphi)^2} = \sqrt{\left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial y}\right)^2 + \left(\frac{\partial\varphi}{\partial z}\right)^2} \tag{4}$$

So the direction cosine of the acoustic ray is expressed as follows:

$$\begin{cases} \frac{d(ncos\alpha)}{ds} = \frac{\partial n}{\partial x} \\ \frac{d(ncos\beta)}{ds} = \frac{\partial n}{\partial y} \\ \frac{d(ncos\gamma)}{ds} = \frac{\partial n}{\partial z} \end{cases} \tag{5}$$

For horizontally layered media, the speed of sound is only a function of the coordinate z, so the right side of the equation above is 0, so it can be derived

$$\frac{cos\alpha}{c(z)} = \frac{cos\alpha_0}{c_0} = C \tag{6}$$

Among them, α is the angle between the sound propagation direction and the horizontal direction, called the glancing angle; $\mathrm{c}(x,y,z) = c(z)$ is the speed of sound at the corresponding depth z. α_0 and c_0 correspond to the value of the acoustic ray, and C is a constant. Equation (6) is Snell's law, which is the basic law of ray acoustics. When $\mathrm{c}(z) > c_0$, $\alpha < \alpha_0$; when $\mathrm{c}(z) < c_0$, $\alpha > \alpha_0$, it is obvious that the sound line always bends to the place where the speed of sound is low.

The intensity equation is used to express the strength of acoustic rays. To define the sound intensity $\mathrm{I} = \mathrm{A}^2\nabla\varphi$, Formula (3) is transformed into $\nabla\cdot\mathrm{I} = 0$, and the volume fraction $\nabla\cdot\mathrm{I}$ is transformed into area fraction by applying the Org-Gao theorem. If the closed surface S is selected as the cross-sections S_1 and S_2 along the side of the acoustic ray tube bundle and both ends of the tube bundle, the area along the side of the acoustic ray tube bundle is zero. Suppose the sound intensity I is uniformly distributed along S_1. and S_2, namely:

$$\mathrm{I}_{S_1}s_1 = \mathrm{I}_{S_2}s_2 = \cdots = con \tag{7}$$

Sound intensity is the sound energy per unit of time that passes through a unit area perpendicular to the direction of sound wave propagation. Let W denote the radiated sound power within a unit solid angle, the basic formula for calculating the sound intensity of a single acoustic ray in ray acoustics:

$$\mathrm{I}(x,z) = \frac{W cos\alpha_0}{r\left|\frac{\partial x}{\partial \alpha_0}\right|_{\alpha_0} sin\alpha_z} \tag{8}$$

where r is the horizontal distance from the launch to the receiving depth. According to the expression of sound intensity I, the expression of sound pressure amplitude can be derived:

$$\mathrm{A}(r,z) = \sqrt{\frac{W cos\alpha_0}{r\left|\frac{\partial x}{\partial \alpha_0}\right|_{\alpha_0} sin\alpha_z}} \tag{9}$$

So the sound field of the ray is expressed as:

$$\mathrm{p}(r,z) = A(r,z)e^{-jk_0\varphi(r,z)} \tag{10}$$

In the formula, $\varphi(x,z) = xcos\alpha_0 + \int \sqrt{n^2 - cos^2\alpha_0}dz + C$ is the deformation of the equation under the premise of setting the boundary conditions, and C is the integral constant.

In particular, the acoustic rays that depart from the sound source and reach the receiver through a certain path are collectively called Eigen acoustic rays, which are the key to the application of ray acoustic theory to describe the sound field. Define $\mathrm{Q} = \frac{cos\alpha_0}{c(z_0)} = \frac{cos\alpha_z}{c(z)}$ as the acoustic ray parameter, then along the Eigen acoustic ray direction:

$$\frac{d^2z}{dx^2} = \frac{Q}{cos^3\alpha_z}\cdot\frac{dc(z)}{dz} = \frac{1}{Q^2c^3(z)}\cdot\frac{dc(z)}{dz} \tag{11}$$

We use the Runge–Kutta method to solve this differential equation. According to the definition of eigenline, the solution $\mathrm{z}(\alpha_0, x)$ should satisfy:

$$\mathrm{z}(\alpha_0, x_r) = z_r \tag{12}$$

By changing the initial glancing angle α_0, using the Newton iteration method to search for and determine the Eigen acoustic rays through the set accuracy range and, then, obtain the acoustic ray propagation time in the moving medium:

$$\mathrm{t} = \int_0^l \frac{ds}{c(z) + u(z)/cos\theta(z)} \tag{13}$$

In the working model of reciprocal transmission, assuming that A and B are the stations of two acoustic signals, the two-way propagation time of the i-th acoustic line can be expressed as:

$$\mathrm{t}^+ = \int_0^{l_+} \frac{ds}{c_0 + \delta c + u/cos\theta} \tag{14}$$

$$\mathrm{t}^- = \int_0^{l_-} \frac{ds}{c_0 + \delta c - u/cos\theta} \tag{15}$$

In the formula, u is the flow velocity of seawater, c_0 is the average speed of sound in the reference environment, and δc represents the amount of disturbance of the speed of sound. t^+ and l_+ are the propagation time and time path of sound waves from station A to station B, t^- and l_- are the propagation time and time path of sound waves from station B to station A. Considering $c_0 \gg \delta c$, only the second-order Taylor expansion is considered, and the higher-order terms in the Taylor expansion are ignored. Then the flow velocity along the sound ray direction can be calculated by the reciprocal propagation time difference, denoted by Δt_i:

$$\begin{aligned}\Delta t_i = t_i^+ - t_i^- &= \int \frac{ds}{c+u/cos\theta} - \int \frac{ds}{c-u/cos\theta} \\ &\approx \sum_{j=1}^{N}\left(\frac{\Delta s^+}{c_0+\delta c} - \frac{\Delta s^+}{(c_0+\delta c)^{+2}\cdot cos\theta}\cdot u\right) \\ &- \sum_{j=1}^{N}\left(\frac{\Delta s^-}{c_0+\delta c} + \frac{\Delta s^-}{(c_0+\delta c)^{-2}\cdot cos\theta}\cdot u\right)\end{aligned} \tag{16}$$

2.2. Flow Field Acoustic Tomography Least Square Method

Equation (16) can be expressed as a matrix form $y = Ex + n$, where y represents the data vector, E is the observation matrix, x is the vector containing the laminar flow unknowns, and n is the noise vector. Under normal circumstances, the number of observations is not equal to the number of tomographic parameters, so a unique solution cannot be obtained by solving linear algebraic equations. Usually, the least square method is used to solve the problem that the number of observations is greater than the number of tomographic parameters, that is, the overdetermined problem and the regularized least square method is used to solve the problem that the number of observations is less than the number of tomographic parameters, that is, the underdetermined problem. The minimized cost function is expressed as:

$$J(x) = \|y - Ex\|^2 + \alpha(x - x_0)^T\gamma(x - x_0) \tag{17}$$

where α represents the regular coefficient, x_0 is the prior estimate of x, and γ represents the regularization matrix. Therefore, the final solution is expressed as:

$$\hat{x} = x_0 + \left(E^T E + \alpha\gamma\right)^{-1} E^T(y - Ex_0) \tag{18}$$

In the process of solving the regularization matrix, γ is selected empirically, and there is no theoretical support, so the application of regularized least squares method has uncertainty. In the actual ocean environment under the Arctic ice, nonlinear interference makes the bandwidth of the received acoustic signal larger [15], resulting in a large error in the time difference obtained by acoustic tomographic reciprocal transmission. Traditional inversion algorithms are more derived from linear theory. In the case of nonlinear disturbances in the inversion estimation, it is easy to produce over-fitting, which has great limitations in the application under Arctic ice. Therefore, an algorithm that can improve the calculation accuracy of acoustic tomography under Arctic ice is urgently needed.

2.3. Flow Field Acoustic Tomography Based on CNN

The application of CNN [16] is a hot topic in the field of acoustics and can play a powerful role in solving inverse problems. For example, Timo et al. [17] (2017) applied convolutional neural network CNN to ultrasonic tomography to solve the problem of porous material parameter estimation, saving labor costs. Araya-Polo et al. [18] (2018) used a deep neural network (DNN) to build a velocity model from seismic traces, which greatly improved the computational efficiency. Compared with the traditional inverse method, the advantage of CNN is not only that it does not need to evaluate the forward model, but also that it can automatically screen important characteristic parameters.

In this study, the acoustic data obtained from the forward model are used as input, and then, the convolution layer is connected to carry out convolution operation on each feature node to extract the characteristic values. Then, pooling processing is carried out to further sample the eigenvalues of the convolutional layer. Using the maximum pooling layer to process data cannot only improve the training efficiency of the convolutional neural network but also improve the robustness and stability of the algorithm. Then, the transpose convolution operation is applied to enlarge the size of the output to be the same as the size of the input. Finally, the flow velocity is predicted and estimated by the nonlinear rectification ReLU function. Figure 2 shows the flow chart of convolutional neural network training:

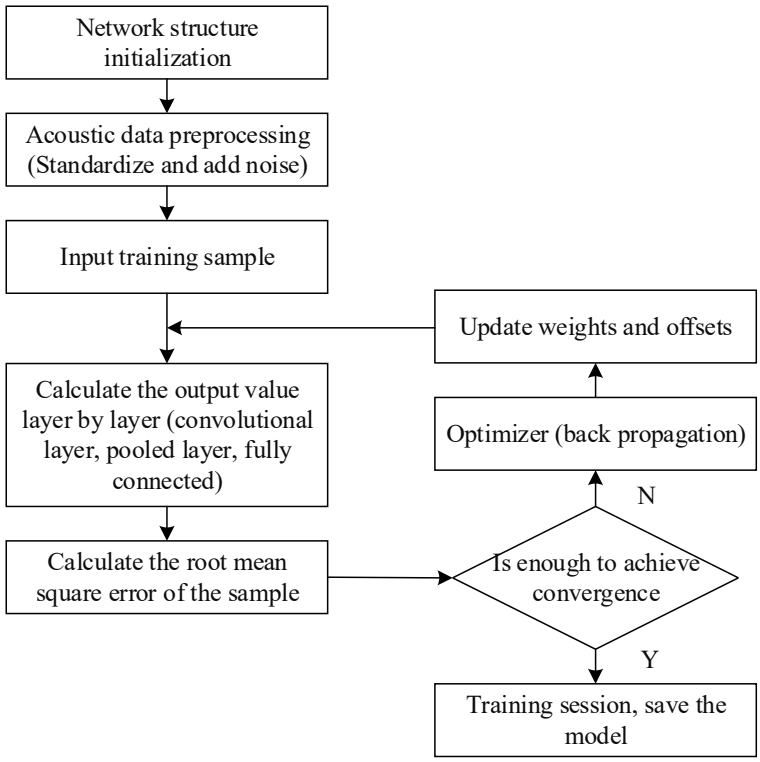

Figure 2. Based on CNN velocity inversion process.

The CNN method is described in mathematical expressions as follows:

$$y = \mathrm{Net}(x;\Theta) = \sigma\left(\omega^{(2)}\cdot\sigma\left(\omega^{(1)}\cdot x + b^{(1)}\right) + b^{(2)}\right) \tag{19}$$

In the formula, Net( ) represents the non-linear mapping of the network, x and y are the input and output of the network, respectively, and Θ is the parameter set that $\left\{\omega^{(1)}, b^{(1)}, \omega^{(2)}, b^{(2)}\right\}$ needs to learn, where $\left(\omega^{(1)}, \omega^{(2)}\right)$ represents the weight between layers, and $\left(b^{(1)}, b^{(2)}\right)$ is the bias. $\sigma(\ )$ is the non-linear rectification function used, including sigmoid, tanh, and ReLU functions. ReLU is a commonly used activation function at present. Its convergence speed is much faster than sigmoid and tanh. It has better performance when building large complex networks and can be given priority.

The application of CNN in Arctic sub-ice acoustic tomography is to find a non-linear mapping function between input and output, learn features from given data, and improve computational efficiency while preventing overfitting. The flow rate inversion estimation process based on CNN is shown in Figure 3. The mathematical expression of the output layer is:

$$u = \mathrm{Net}(a;\Theta) \tag{20}$$

Among them, a is the original acoustic dataset, including the forward and backward propagation time, incident angle, sound velocity, trajectory length, and other characteristics of the ray, and u represents the flow velocity predicted by the network. The application of convolutional neural network requires two processes of training and prediction. Before training the dataset, first, establish a forward model and combine the ray theory equation to simulate the sound propagation process in a moving medium, to obtain acoustic data as input. Input the input and the corresponding output into the network for training and learning and continuously optimize and adjust the weight and bias during the training process to obtain the nonlinear mapping relationship from input to output. The optimization benchmark is to find a set of weights and biases that minimizes the difference between

the output and the predicted value given by the network. Therefore, the loss function f is expressed as:

$$\mathrm{f} = \mathrm{argmin}\frac{1}{mN}\sum_{n=1}^{N} L(Net(a;\Theta),u) \tag{21}$$

where m is the total number of samples, and $L()$ represents the error between the predicted value and the true value. Then, choose L_2 regularization method to improve generalization ability, leave the most relevant features and reduce the weight of unimportant features.

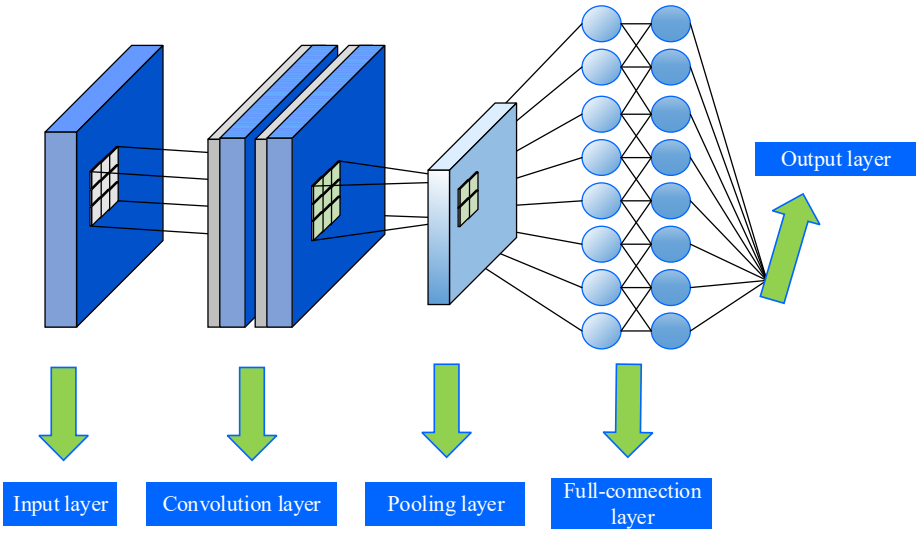

Figure 3. CNN model construction process.

Use back propagation and gradient descent to update the learning parameter Θ. Using the chain derivation rule to obtain the network model gradient layer by layer from back to front, by randomly selecting a small batch of samples each time, the network model parameters are continuously updated along the negative direction of the gradient (first derivative) of the loss function. Until the change in the gradient reaches the specified threshold or no more changes, the learning process is stopped. The optimization problem is expressed as:

$$\mathrm{f} = \mathrm{argmin}\frac{1}{mK}\sum_{n=1}^{K} \|u - Net(a;\Theta)\|_2^2 \tag{22}$$

The network parameter update expression is as follows:

$$\Theta_{t+1} = \Theta_t - \varepsilon\frac{1}{m}\nabla_{\Theta}\sum_{n=1}^{m} L(a_t;u_t;\Theta_t) \tag{23}$$

Among them, the number of training steps is represented by t, the learning rate is ε, the amount of data in each batch is represented by m during small-batch learning, the input value of the nth sample of each batch of data is a_t, and the target true value is u_t. The loss function of the predicted value and the true value is denoted by $L(\cdot)$. After the model training is completed, it enters the prediction stage and inputs new acoustic data to the saved network to obtain the flow velocity in a certain area.

3. Modeling and Simulation

3.1. Model Establishment

Establish a marine test environment [19], the density of seawater is 1000 kg/m^3, the speed of sound on the sea surface is 1530 m/s, the speed of seafloor is 1500 m/s, and there is a certain sound velocity gradient. The seawater and the seabed are regarded as the fluid-fluid interface, and the sound velocity of the adjacent fluid is set to 1575 m/s, the density

is 1700 kg/m^3, and the attenuation coefficient is changed by $0.9 \times f_0/(20 \times c_1 \times lge)$. The sea depth is 110 m, the horizontal range is 1500 m, the sea surface is set as the Rayleigh roughness model, and the root mean square roughness is 1 m. The three sound sources are at a depth of 25 m, 35 m, and 45 m, and the total source power of the emitted sound waves is 1 [W/m]. For all calculation domains including seawater and seabed, free triangular meshes are used for subdivision. Considering the computational complexity and efficiency issues, the maximum unit size customized for this study is 20 m. The time step set by the study is range (0, 0.001, 1) s. Figures 4–7 show the ray propagation at each time point.

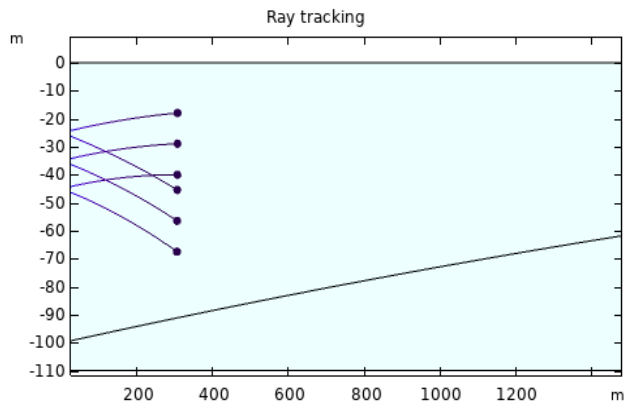

Figure 4. Ray trajectories at t = 0.2 s.

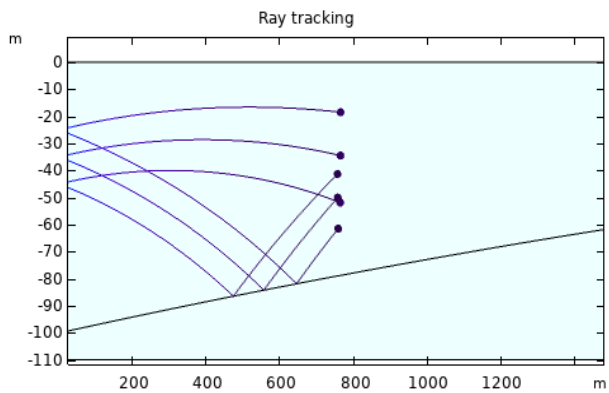

Figure 5. Ray trajectories at t = 0.5 s.

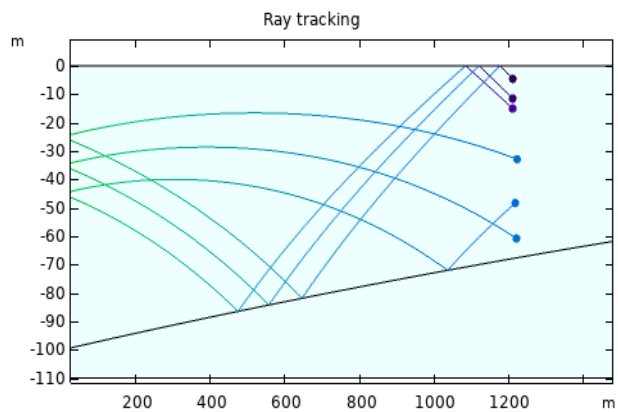

Figure 6. Ray trajectories at t = 0.8 s.

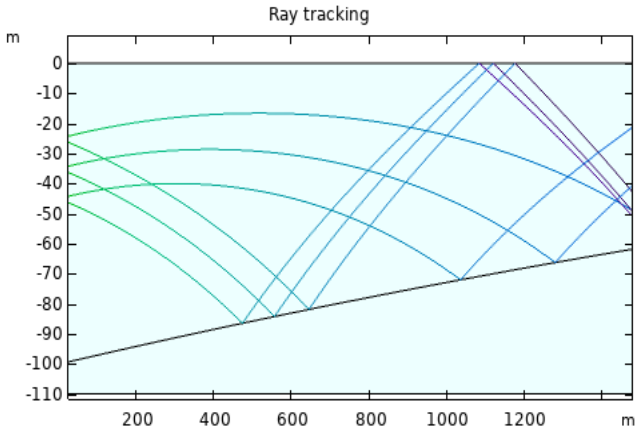

Figure 7. Ray trajectories at t = 1.0 s.

Assuming that the seawater velocity is a one-dimensional velocity structure, the model is divided into 3 layers, the velocity of 0–20 m is 2 m/s, the velocity of 20–40 m is 1 m/s, and the velocity of 40–110 m is 0.5 m/s. The eastward flow rate is positive, and the westward flow rate is negative. Obtain the acoustic ray's forward and reverse propagation time as shown in Table 1.

Table 1. Forward and backward propagation time.

	Forward Propagation Time/s	Backward Propagation Time/s	△t/s
1	0.9908667974750063	0.9921370357778033	−0.001270238
2	0.9814290042702533	0.9832276442346188	−0.00179864
3	0.9911920895784465	0.9924207624924704	−0.001228673
4	0.9845382912123009	0.9855862764180756	−0.001047985
5	0.9909932804115785	0.9922227519661865	−0.001229472
6	0.9861684623886157	0.9869143234892909	−0.000745861

3.2. Flow Field Acoustic Tomography Error and Accuracy Judgment Index

In order to evaluate the accuracy of the least squares algorithm and the convolutional neural network inverting the flow velocity, this paper measures the accuracy of the prediction model based on the average absolute percentage error, an evaluation index. Calculated as follows:

$$\text{MAPE} = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{y_i - u_i}{y_i}\right| \times 100\% \tag{24}$$

In the formula, y_i represents the real flow velocity value, and u_i is the predicted flow velocity value. The average absolute percentage error indicates the degree of deviation between the predicted value of the flow rate and the actual value. It is more reasonable to use this indicator to evaluate the flow rate prediction effect under the level stratification. At the same time, the condition number is introduced to judge the stability of the matrix $\Delta t = Eu + n$ solved by the least squares, and the condition number of the matrix E is used to measure the sensitivity of the output of the matrix multiplication inverse to the input error. If Δt is slightly disturbed, the condition number of E is larger, and the flow velocity u obtained at this time will have a big change. In other words, the larger the condition number, the worse the sensitivity.

4. Simulation Results

4.1. Least Squares Simulation Results

Based on the least squares flow velocity tomography method, for a simple flow field structure, the characteristics of the flow field can be basically reconstructed. Table 2 shows the calculated values obtained by releasing different numbers of acoustic rays and the corresponding error results.

According to the tomographic results in Figure 8, the least squares method can be used to chromatogram the structural features of a simple flow field, u0 represents the true value, and u11 is the predicted value. As can be seen from Table 2, when the number of velocity layers is three, the velocity result obtained by releasing three acoustic rays is closer to the true value than that obtained by releasing more than three acoustic rays, which indicates that the essence of the least square method does not obtain the accurate solution but finds the optimal solution. At the same time, according to the average absolute percentage error (MAPE) value in the table, the inversion estimation accuracy of flow velocity of 1 m/s is higher than that of flow velocity of 0.5 m/s, indicating that the larger the flow velocity is, the more accurate the inversion results will be. It can also be found from the table that when six acoustic rays are released, the condition number of the matrix calculated is 9.0633, and as the acoustic ray increase, the matrix condition number also increases, indicating that the perturbation of time difference will have a great impact on the result. That is, when simulating an uneven ocean environment, the inversion result based on the regularized least square tomographic flow velocity will be inaccurate.

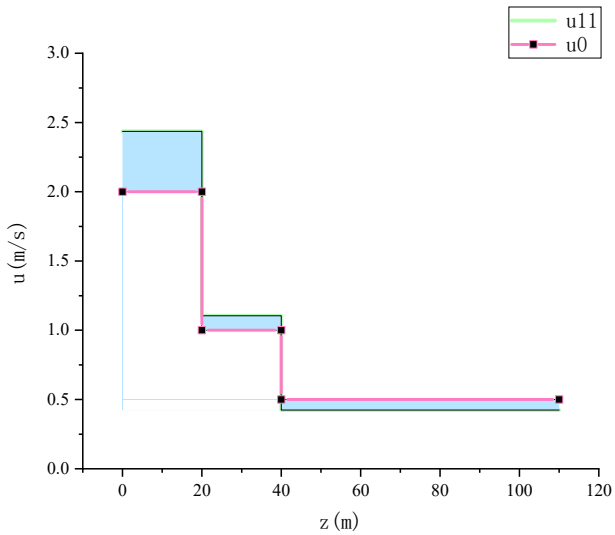

Figure 8. Least squares method to obtain the velocity compared with the real value.

Table 2. Flow velocity inversion results.

	0–20	20–40	40–110	Condition Number	Mean Absolute Error
3 MAPE	2.43632 0.21816	1.10469 0.10469	0.42396 0.15208	6.07417	0.15831
4 MAPE	2.24261 0.121305	1.01349 0.01349	0.77710 0.5542	6.69373	0.229665
5 MAPE	2.74908 0.37454	0.68197 0.31803	0.75077 0.50154	7.74111	0.398037
6 MAPE	2.5162 0.2581	0.8419 0.1581	0.7991 0.5982	9.0633	0.338133

4.2. Implementation of CNN

The implementation of CNN is based on the Keras framework and is built using a sequential model. According to the parameters mentioned in Equation (16) and the simulation of 3.1 forward model, seven characteristic values of incidence angle, depth, sound velocity, and propagation time difference of acoustic ray in forward and reverse time flow are obtained, respectively, and 168 groups of original acoustic data are extracted as input. The output data are the unique flow rate value corresponding to each group of input data. In order to make the data better as the input of the convolutional neural network, the

original data are also standardized to make the entire neural network more stable in the middle output value of each layer.

Under supervised learning, the test dataset (size = 0.3) relies on the optimal model obtained in advance training and uses this model to map the input to obtain the corresponding output, so as to achieve the purpose of flow velocity inversion. The specific implementation process is shown in Figure 9.

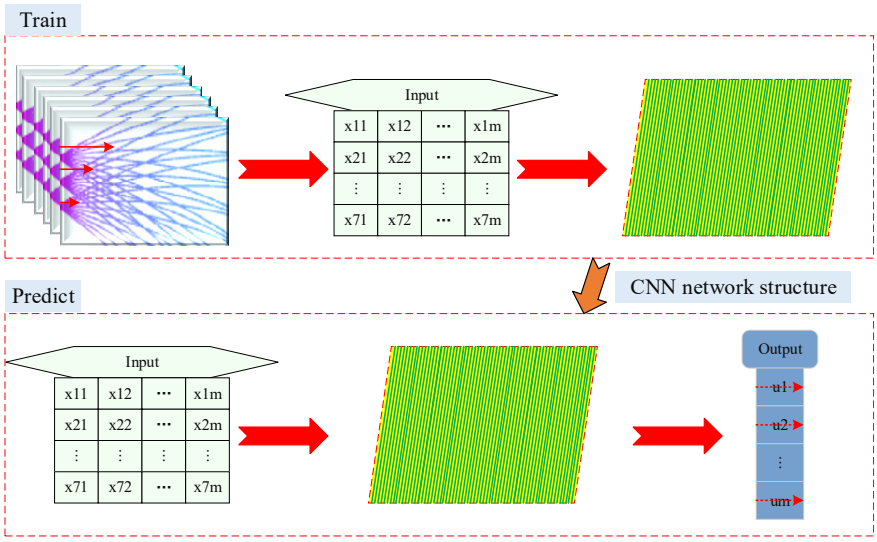

Figure 9. Convolutional neural network training flow chart.

The design of the convolutional neural network uses four convolutional layers, namely, Conv1, Conv2, Conv3, and Conv4. After each layer of convolution operation, the ReLU activation function is used for correction, and then, the sampling layer is used for the pooling calculation. The size of the convolution kernel is 3×3, and a pooling layer is added after every two convolution layers. The pooling size of each layer is 2×2, and the maximum value in the pooling area is taken as the neuron output value of the pooling layer, so as to improve the calculation efficiency. After the second pooling layer, connect the fully connected layer, expand the pooling result obtained in the previous step into a one-dimensional feature vector, adjust the parameter Dropout = 0.5 to prevent overfitting, and finally, obtain the flow velocity data through the linear function output.

Different from the use of the convolutional neural network for image recognition, this study uses one-dimensional [19] data, and the number of network channels is no longer three but one. Before performing the convolution operation, use valid padding to fill the boundary of the matrix to increase the size of the matrix so that the input and output sizes are the same.

4.3. Experimental Analysis of CNN

In the process of model training, the Adam optimizer is selected to verify the accuracy of the model, and the mean square error is used as the loss function. The convergence of the model is judged by changing the number of iterations. When the loss value gradually approaches a certain value and only fluctuates in a small range, it is considered that the convergence state has been reached. After setting the hyperparameters, the network that is considered to be more effective is selected. Figure 10 is the loss curve obtained by making the model iterate 500 times.

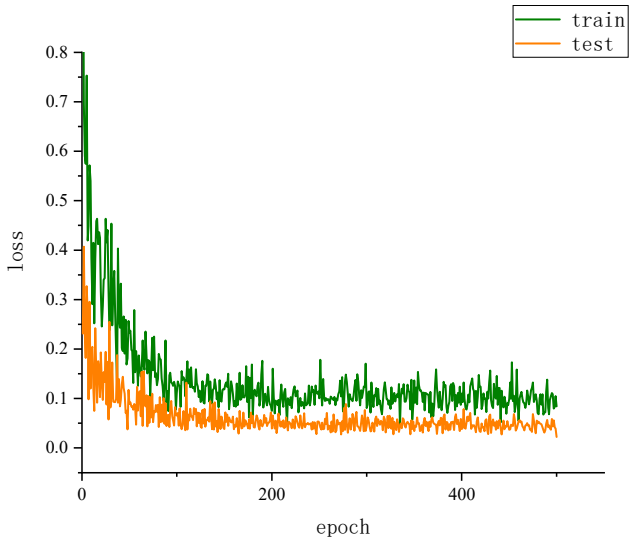

Figure 10. Training and testing of number of epochs.

Figure 11 shows an exemplary result of applying this method. The results show that the 0–20 m depth velocity value is 1.9283856 m/s, the 20–40 m velocity value is 1.0018886 m/s, and the 40–110 m velocity value is estimated to be 0.5931605 m/s. It can be seen from the results that there is a good match between the predicted value and the true value. Based on this, the accuracy of the prediction model is further measured, and the calculated average absolute percentage error is 0.074672267, which is less than the average percentage error obtained by the least squares. Therefore, in a uniform ocean environment, using CNN to perform flow velocity inversion estimation results is more accurate.

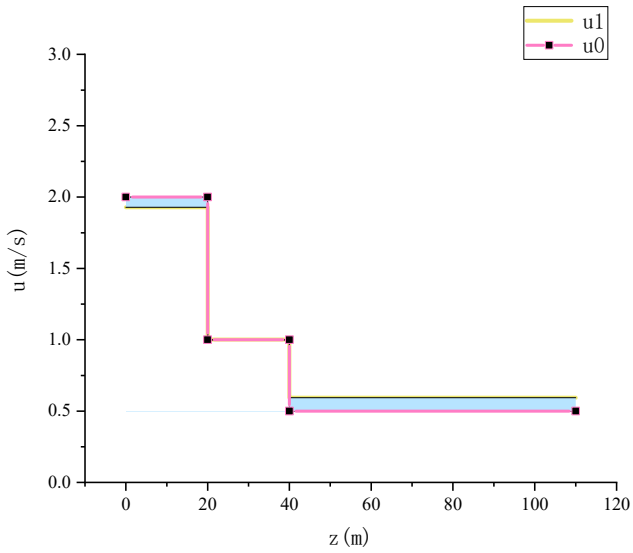

Figure 11. CNN to obtain the velocity compared with the real value.

In order to further verify the effectiveness of the proposed method, we add Gaussian white noise with a zero mean standard deviation of $5 \times e^{-8}$ to each set of time difference feature parameters [20,21] to simulate the inhomogeneity of the marine environment. Figure 12 shows the predicted results using the proposed method with noisy input. The velocity value of 0–20 m depth is 1.9210289 m/s, the velocity value of 20–40 m is 1.0035175 m/s, and the estimated velocity of 40–110 m is 0.6070473 m/s. The calculated average absolute percentage error is 0.085699217.

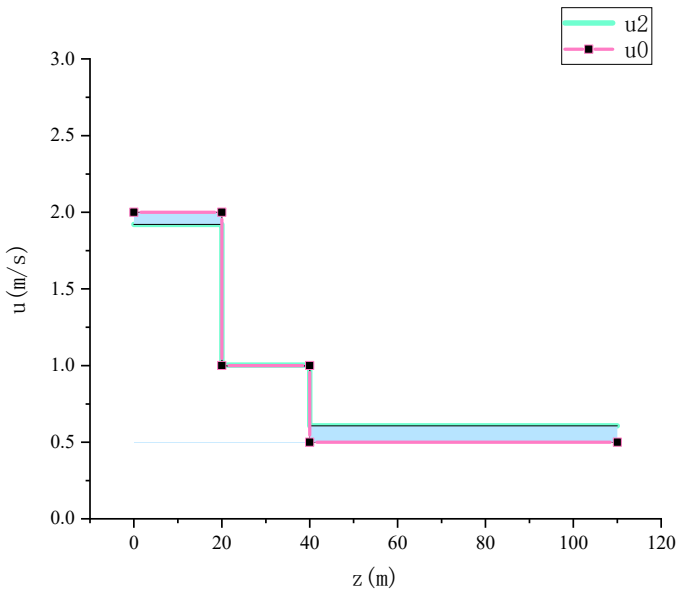

Figure 12. CNN to obtain the velocity compared with the real value under disturbance.

Compared with the case of no interference, the prediction result with noisy input is slightly worse, but it is still better than the traditional inversion method, revealing that the proposed method can still produce acceptable results in the case of disturbance. All in all, even if it is disturbed by the nonlinearity of the actual marine environment, the convolutional neural network can estimate the actual flow velocity value through nonlinear mapping inversion. In addition, this method has high calculation efficiency. It only needs to consume time cost in the training model stage. After the optimal model is saved, the corresponding velocity value can be obtained by only inputting each index parameter in the prediction work. Therefore, it is proved that the inversion algorithm based on the convolutional neural network can stably realize the feasibility and reliability of the flow rate output.

5. Conclusions

This paper mainly focuses on the study of acoustic tomography methods for the flow field under ice in polar regions. Under the cover of Arctic sea ice, sub-ice acoustic tomography is the only feasible method for large-scale current measurement in the Arctic. In high latitude areas, under the action of Coriolis force, the unevenness of the marine environment greatly affects the accuracy of sub-ice acoustic tomography. Through theoretical research and simulation of the flow field tomography of the regularized least squares method, this kind of disturbance cannot be characterized in the calculation of sub-ice acoustic tomography in a constant form or an empirical function, and it is difficult to continue to be applied in the case of uneven seawater. Based on this, an acoustic tomographic ocean current inversion estimation method based on convolutional neural network is proposed.

This seawater disturbance is introduced into the acoustic tomography calculation in the form of a Gaussian function. With the help of COMSOL and programming software, the finite element method is used to construct the sound field and marine environmental parameters. The results show that CNN can effectively approximate the inverse of the nonlinear operator. When in a real uneven ocean environment, the constructed network can still calculate a satisfactory flow velocity value. Moreover, this method has high computational efficiency and only requires time and cost in the training model stage. After the optimal model is saved, only the various index parameters need to be input when performing prediction work to obtain the corresponding flow rate value. The time in the entire forecasting process is negligible. The simulation results verify that CNN can stably

realize the effectiveness and reliability of flow velocity reconstruction in the under-ice acoustic tomography problem.

Author Contributions: Conceptualization, K.J., J.X. and Z.W.; methodology, Z.L. and L.F.; software, K.J., C.L. and Z.W.; validation, K.J., J.X., Z.W., C.L., L.F., Z.L. and J.Z.; formal analysis, K.J., J.X., Z.W., C.L., L.F., Z.L. and J.Z.; investigation, K.J., J.X., Z.W., C.L., L.F., Z.L. and J.Z.; resources, L.F., Z.L. and J.Z.; writing—original draft preparation, K.J.; writing—review and editing, J.X.; visualization, Z.W.; supervision, C.L.; project administration, L.F., Z.L. and J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was jointly supported by the National Natural Science Foundation of China (41706106).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Maeda, K.; Kimura, N.; Yamaguchi, H. Temporal and spatial change in the relationship between sea-ice motion and wind in the Arctic. *Polar Res.* **2020**, *39*. [CrossRef]
2. Wunsch, M.C. Ocean acoustic tomography: A scheme for large scale monitoring. *Deep Sea Res.* **1979**. [CrossRef]
3. Zhang, Y.; Maslowski, W.; Semtner, A.J. Impact of mesoscale ocean currents on sea ice in high-resolution Arctic ice and ocean simulations. *J. Geophys. Res. Oceans* **1999**, *104*, 18409–18429. [CrossRef]
4. Spindel, R.C.; Worcester, P.F. Ocean Acoustic Tomography. *Sci. Am.* **1990**, *263*, 94–99. [CrossRef]
5. Skarsoulis, E.K. A matched-peak inversion approach for ocean acoustic travel-time tomography. *J. Acoust. Soc. Am.* **2000**, *107*, 1324–1332. [CrossRef] [PubMed]
6. Taniguchi, N.; Huang, C.-F.; Kaneko, A.; Liu, C.-T.; Howe, B.M.; Wang, Y.-H.; Yang, Y.; Lin, J.; Zhu, X.-H.; Gohda, N. Measuring the Kuroshio Current with ocean acoustic tomography. *J. Acoust. Soc. Am.* **2013**, *134*, 3272–3281. [CrossRef] [PubMed]
7. Jin, G.; Pan, H.; Zhang, Q.; Lv, X.; Zhao, W.; Gao, Y. Determination of Harmonic Parameters with Temporal Variations: An Enhanced Harmonic Analysis Algorithm and Application to Internal Tidal Currents in the South China Sea. *J. Atmos. Ocean. Technol.* **2018**, *35*. [CrossRef]
8. Dushaw, B.D.; Sagen, H. The role of simulated small-scale ocean variability in inverse computations for ocean acoustic tomography. *J. Acoust. Soc. Am.* **2017**, *142*, 3541. [CrossRef] [PubMed]
9. Dushaw, B.D.; Sagen, H. On the effects of small-scale variability on acoustic propagation in Fram Strait: The tomography forward problem. *J. Acoust. Soc. Am.* **2016**, *140*, 1286. [CrossRef] [PubMed]
10. Callaham, J.; Maeda, K.; Brunton, S.L. Robust flow field reconstruction from limited measurements via sparse representation. *Phys. Rev. Fluids* **2019**, *4*, 103907. [CrossRef]
11. Mustaqeem; Kwon, S. 1D-CNN: Speech Emotion Recognition System Using a Stacked Network with Dilated CNN Features. *Comput. Mater. Con.* **2021**, *67*, 4039–4059. [CrossRef]
12. Stéphan, Y.; Thiria, S.; Badran, F. Application of multilayered neural networks to ocean acoustic tomography inversions. *Inverse Probl. Eng.* **1996**, *3*, 281–304. [CrossRef]
13. Narvekar, C.; Rao, M. Flower classification using CNN and transfer learning in CNN- Agriculture Perspetive. In Proceedings of the 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, 3–5 December 2020. [CrossRef]
14. Powell, B.S.; Kerry, C.G.; Cornuelle, B.D. Using a numerical model to understand the connection between the ocean and acoustic travel-time measurements. *J. Acoust. Soc. Am.* **2013**, *134*, 3211–3222. [CrossRef] [PubMed]
15. Skarsoulis, E.; Piperakis, G.; Kalogerakis, M.; Sagen, H. Ocean acoustic tomography: Travel-time inversion in the eastern Fram Strait. *J. Acoust. Soc. Am.* **2008**, *123*, 2991. [CrossRef]
16. Frederick, C.; Villar, S.; Michalopoulou, Z.H. Seabed classification using physics-based modeling and machine learning. *J. Acoust. Soc. Am.* **2020**, *148*, 859–872. [CrossRef] [PubMed]
17. Lähivaara, T.; Kärkkäinen, L.; Huttunen, J.; Hesthaven, J.S. Deep convolutional neural networks for estimating porous material parameters with ultrasound tomography. *J. Acoust. Soc. Am.* **2018**, *143*, 1148. [CrossRef] [PubMed]
18. Yang, F.; Ma, J. Deep-learning inversion: A next generation seismic velocity-model building method. *Geophysics* **2019**, 1–133. [CrossRef]
19. Etter, P.C. *Underwater Acoustic Modeling and Simulation*; Spon Press (Tay & Francis Group): London, UK, 2003. [CrossRef]

20. Comanducci, L.; Bestagini, P.; Tagliasacchi, M.; Sarti, A.; Tubaro, S. Reconstructing Speech from CNN Embeddings. *IEEE. Signal. Proc. Let.* **2021**, *28*, 952–956. [CrossRef]
21. Blondel, P.; Sagen, H.; Martin, B.; Pettit, E.; Tegowski, J.; Thodes, A.; Tollefsen, D.; Worcester, P. Report of the Polar Session, Oceanoise2015: Report of the Round Table Session on Polar Noise. Available online: http://oceanoise2015.com/?page_id=789 (accessed on 6 July 2021).